LETTER

https://doi.org/10.1038/s41586-019-1043-4

Complex societies precede moralizing gods throughout world history

Harvey Whitehouse[1,15], Pieter François[1,2,15], Patrick E. Savage[1,3,15]*, Thomas E. Currie[4], Kevin C. Feeney[5], Enrico Cioni[6], Rosalind Purcell[6], Robert M. Ross[1,7,8], Jennifer Larson[9], John Baines[10], Barend ter Haar[11], Alan Covey[12] & Peter Turchin[13,14]

The origins of religion and of complex societies represent evolutionary puzzles[1–8]. The 'moralizing gods' hypothesis offers a solution to both puzzles by proposing that belief in morally concerned supernatural agents culturally evolved to facilitate cooperation among strangers in large-scale societies[9–13]. Although previous research has suggested an association between the presence of moralizing gods and social complexity[3,6,7,9–18], the relationship between the two is disputed[9–13,19–24], and attempts to establish causality have been hampered by limitations in the availability of detailed global longitudinal data. To overcome these limitations, here we systematically coded records from 414 societies that span the past 10,000 years from 30 regions around the world, using 51 measures of social complexity and 4 measures of supernatural enforcement of morality. Our analyses not only confirm the association between moralizing gods and social complexity, but also reveal that moralizing gods follow—rather than precede—large increases in social complexity. Contrary to previous predictions[9,12,16,18], powerful moralizing 'big gods' and prosocial supernatural punishment tend to appear only after the emergence of 'megasocieties' with populations of more than around one million people. Moralizing gods are not a prerequisite for the evolution of social complexity, but they may help to sustain and expand complex multi-ethnic empires after they have become established. By contrast, rituals that facilitate the standardization of religious traditions across large populations[25,26] generally precede the appearance of moralizing gods. This suggests that ritual practices were more important than the particular content of religious belief to the initial rise of social complexity.

Supernatural agents that punish direct affronts to themselves (for example, failure to perform sacrifices or observe taboos) are commonly represented in global history, but rarely are such deities believed to punish moral violations in interactions between humans[2]. Recent millennia, however, have seen the rise and spread of several 'prosocial religions', which include either powerful 'moralizing high gods' (MHG; for example, the Abrahamic God) or more general 'broad supernatural punishment' (BSP) of moral transgressions (for example, karma in Buddhism)[9,12,16–18]. Such moralizing gods may have provided a crucial mechanism for overcoming the classic free-rider problem in large-scale societies[11]. The association between moralizing gods and complex societies has been supported by two forms of evidence: psychological experiments[3,6,27,28] and cross-cultural comparative analyses[7,11,14–18,20].

The contributions of theistic beliefs to cooperation, as well as the historical question of whether moralizing gods precede or follow the establishment of large-scale cooperation, have been much debated[9,10,12,23,24]. Three recent studies that explicitly model temporal causality have come to contrasting conclusions. One study, which applied phylogenetic comparative methods to infer historical changes in Austronesian religions, reported that moralizing gods (BSP but not MHG) preceded the evolution of complex societies[16]. The same conclusion was reached in an analysis of historical and archaeological data from Viking-age Scandinavia[18]. By contrast, another study of Eurasian empires has reported that moralizing gods followed—rather than preceded—the rise of complex, affluent societies[20]. However, all of these studies are restricted in geographical scope and use proxies for social complexity that the authors themselves concede are 'very crude'[20] (for example, the binary classification of societies as of either high or low complexity).

To overcome these limitations, we used 'Seshat: Global History Databank'[29], a repository of standardized data on social structure, religion and other domains for hundreds of societies throughout world history. In contrast to other databases that attempt to model history using contemporary ethnographic data, Seshat directly samples over time as well as space. Seshat also includes estimates of expert disagreement and uncertainty, and uses more-detailed variables than many databases.

To test the moralizing gods hypothesis, we coded data on 55 variables from 414 polities (independent political units) that occupied 30 geographical regions from the beginning of the Neolithic period to the beginning of Industrial and/or colonial periods (Fig. 1 and Supplementary Data). We used a recently developed and validated measure of social complexity that condenses 51 social complexity variables (Extended Data Table 5) into a single principal component that captures three quarters of the observed variation, which we call 'social complexity'[8]. The remaining four variables were selected to test the MHG and BSP subtypes of the moralizing gods hypothesis. The MHG variable was coded following the MHG variable used as standard in the literature on this topic[11,14–17,30], which requires that a high god who created and/or governs the cosmos actively enforces human morality. Because the concept of morality is complex, multidimensional and in some respects culturally relative—and because not all moralizing gods are 'high gods'—we also coded three different variables related to BSP that are specifically relevant to prosocial cooperation: reciprocity, fairness and in-group loyalty. For analysis, these three variables were combined into a single BSP variable. The Methods, Supplementary Information and http://seshatdatabank.info/methods/codebook provide further methodological details, definitions and justifications, including a discussion of the relationship between MHG, BSP and big gods.

Figure 1 and Extended Data Table 1 show the temporal and geographical distribution of the appearance of moralizing gods in our sample. Although societies in all 30 regions possessed beliefs about appeasing supernatural agents through the performance of rituals, in 10 out of the 30 regions, there was no evidence for moralizing gods before their introduction under colonial powers. The remaining 20 regions

[1]Centre for the Study of Social Cohesion, University of Oxford, Oxford, UK. [2]St Benet's Hall, Oxford, UK. [3]Faculty of Environment and Information Studies, Keio University, Fujisawa, Japan. [4]Human Behaviour & Cultural Evolution Group, Department of Biosciences, University of Exeter, Penryn, UK. [5]School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland. [6]Seshat: Global History Databank, Evolution Institute, San Antonio, FL, USA. [7]ARC Centre of Excellence in Cognition and its Disorders and Department of Psychology, Royal Holloway, University of London, Egham, UK. [8]Department of Anthropology and Archaeology, University of Bristol, Bristol, UK. [9]Department of Modern & Classical Language Studies, Kent State University, Kent, OH, USA. [10]Faculty of Oriental Studies, University of Oxford, Oxford, UK. [11]Department of Chinese Language and Culture, Asia-Africa-Institute, University of Hamburg, Hamburg, Germany. [12]Department of Anthropology, University of Texas at Austin, Austin, TX, USA. [13]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA. [14]Complexity Science Hub Vienna, Wien, Austria. [15]These authors contributed equally: Harvey Whitehouse, Pieter François, Patrick E. Savage. *e-mail: psavage@sfc.keio.ac.jp

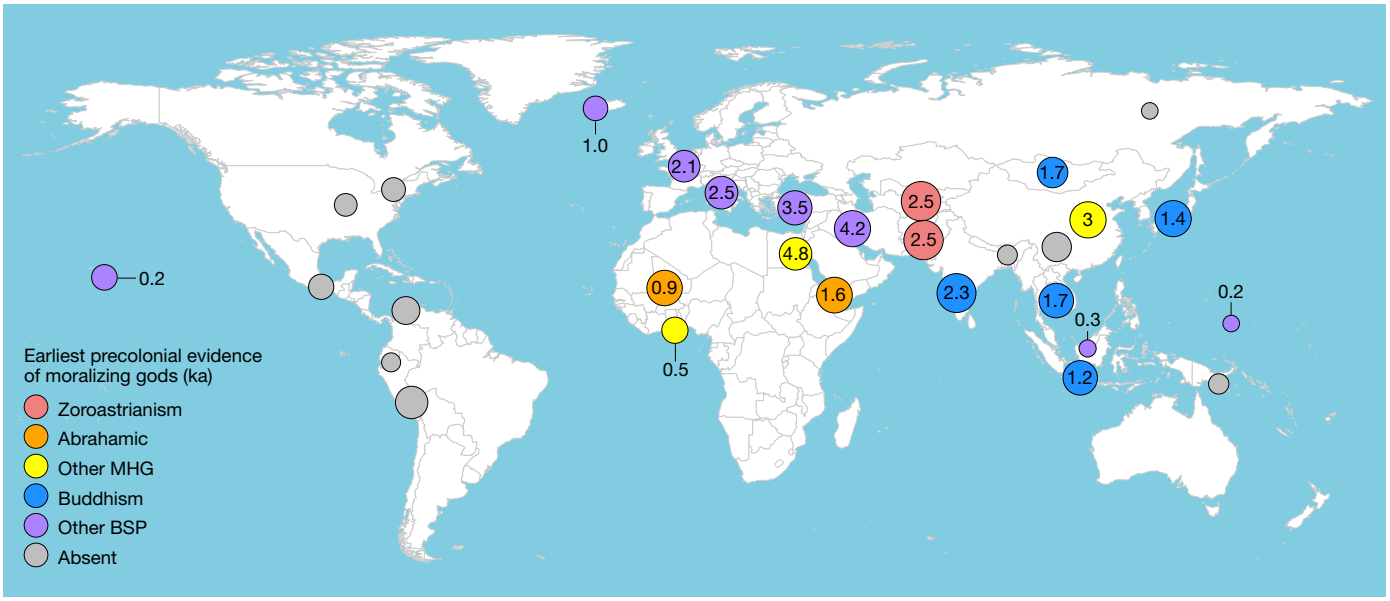


Fig. 1 | Locations of the 30 sampled regions on the world map, labelled according to precolonial evidence of moralizing gods. The area of each circle is proportional to social complexity of the earliest polity with moralizing gods to occupy the region or the latest precolonial polity for regions without precolonial moralizing gods. For regions with precolonial moralizing gods, the date of earliest evidence of such beliefs is displayed in thousands of years ago (ka), coloured by type of moralizing gods. The three transnational religious systems that represent the first appearance of moralizing gods in more than one region—Zoroastrianism, Abrahamic religions (Judaism, Islam and Christianity) and Buddhism—are coloured red, orange and blue, respectively, whereas other local religious systems with beliefs in MHG or BSP are coloured yellow and purple, respectively. See Extended Data Table 1 for further details.

displayed a diverse range of 15 different systems of belief in moralizing gods: in some, the first evidence of moralizing gods came in the form of MHG and in others it came in the form of BSP (Extended Data Table 1). The first appearance of moralizing gods in our sample was in Egypt, where the concept of supernatural enforcement of *Maat* (order) is attested by the Second Dynasty, around 2800 BC. This was followed by sporadic appearances in local religions throughout Eurasia (Mesopotamia (around 2200 BC), Anatolia (around 1500 BC) and China (around 1000 BC)) before the wider spread of transnational religions began during the first millennium BC with Zoroastrianism and Buddhism, followed later by Christianity and Islam. Although Christianity and Islam would eventually become the most widespread religions, local forms of moralizing gods were present well before they arrived in most regions (for example, Roman gods were believed to punish oath-breaking from as early as 500 BC, almost a millennium before Christianity was adopted as the official Roman religion). The diverse range of religious systems represented in our global sample makes it possible to draw more general conclusions about religion than have previously been possible.

Although our sampling scheme reduces non-independence, our polities still cannot be considered statistically independent because of the historical relationships among them. We controlled for these using a logistic regression model to account for temporal, geographical and cultural dependencies in the global distribution of moralizing gods (see Methods). This analysis revealed that social complexity was a stronger predictor of moralizing gods than temporal, geographical or linguistic relationships, and remained highly significant even after controlling for these relationships ($z = 6.8$, degrees of freedom (d.f.) = 800, $P < 1 \times 10^{-11}$; Extended Data Table 2), conceptually replicating previous studies[7,11,14,15].

The moralizing gods hypothesis posits a 'statistical causal relationship'[10] in which moralizing gods facilitate the evolution of complex societies[9,12,16–18]. This indicates that, on average, social complexity should increase more rapidly following the appearance of moralizing gods. To test this prediction, we conducted time-series analyses of the 12 regions for which social complexity data were available both before and after the appearance of moralizing gods (Fig. 2, Extended Data Table 1 and Extended Data Fig. 1). Notably, average rates of increase of social complexity were over five times greater before—not after—the appearance of moralizing gods (paired *t*-test, $t = -6.6$, d.f. = 199, $P < 1 \times 10^{-9}$; Fig. 2). This trend was significant both globally and individually for 10 out of the 12 regional time-series analyses (Extended Data Table 1 and Extended Data Fig. 1). None of these 12 regions displayed a significantly greater rate of increase in social complexity after the appearance of moralizing gods than before. Robustness analyses showed that our primary finding of higher rates of increasing social complexity before the appearance of moralizing gods was present regardless of the type of moralizing gods (MHG or BSP), the choice of variables used to estimate social complexity, uncertainty in the timing of appearance of moralizing gods, or the time windows used to estimate rates of change in social complexity (Extended Data Table 4).

In summary, although our analyses are consistent with previous studies that show an association between moralizing gods and complex societies[7,11,14–18,30], we find that moralizing gods usually follow—rather than precede—the rise of social complexity. Notably, most societies that exceeded a certain social complexity threshold developed a conception of moralizing gods. Specifically, in 10 out of the 12 regions analysed, the transition to moralizing gods came within 100 years of exceeding a social complexity value of 0.6 (which we call a megasociety, as it corresponds roughly to a population in the order of one million; Extended Data Fig. 1). This megasociety threshold does not seem to correspond to the point at which societies develop writing, which might have suggested that moralizing gods were present earlier but were not preserved archaeologically. Although we cannot rule out this possibility, the fact that written records preceded the development of moralizing gods in 9 out of the 12 regions analysed (by an average period of 400 years; Supplementary Table 2)—combined with the fact that evidence for moralizing gods is lacking in the majority of non-literate societies[2]—suggests that such beliefs were not widespread before the invention of writing. The few small-scale societies that did display precolonial evidence of moralizing gods came from regions that had previously been used to support the claim that moralizing gods contributed to the rise of social complexity (Austronesia[16] and Iceland[18]), which suggests that such regions are the exception rather than the rule.

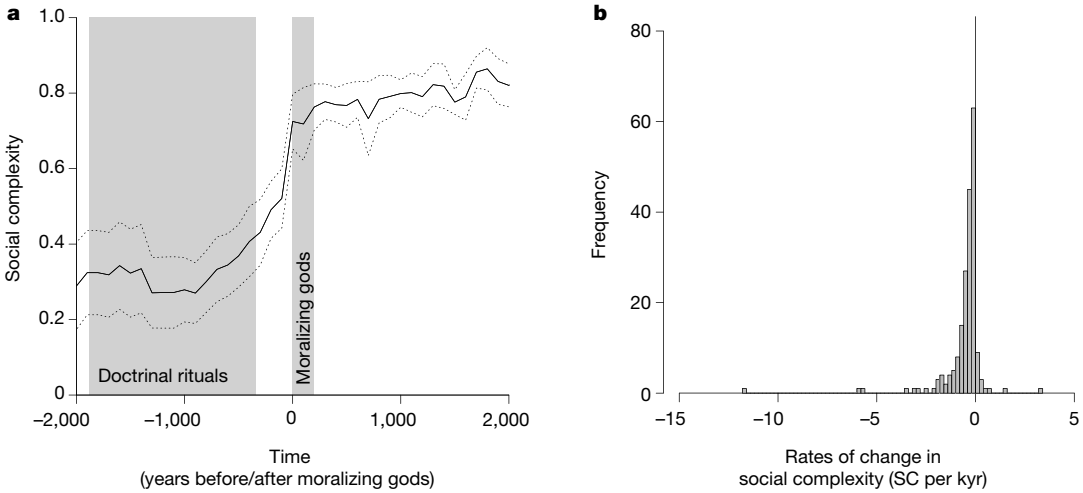


Fig. 2 | Social complexity before and after the appearance of moralizing gods. a, Time series showing mean social complexity over time for 2,000 years before and after the appearance of moralizing gods. $n = 12$ regions with social complexity data for before and after moralizing gods. Social complexity has been scaled so that the society with the highest social complexity (Qing Dynasty, China, around AD 1900) has a value of 1 and the society with the lowest social complexity (Early Woodland, Illinois, USA, around 400 BC) has a value of 0. Vertical bands represent the period in which moralizing gods and doctrinal rituals first appeared. All errors represent 95% confidence intervals, with the exception of the vertical bar for moralizing gods, which represents the mean duration of the polity in which moralizing gods appeared (because times are normalized to the time of first evidence of moralizing gods, and there is thus no variance in this parameter). **b**, Histogram of the differences in rates of change in social complexity (SC) after minus before the appearance of moralizing gods. $n = 200$ time windows from the 12 regions. kyr, thousand years. The y axis represents the number of time windows out of 200. See Extended Data Fig. 1 for data for each of the 12 regions and Extended Data Fig. 2 for a version extending beyond 2,000 years before and after moralizing gods. The analyses in this figure treat the presence of either MHG or BSP as 'moralizing gods'—see Extended Data Fig. 3 for an alternative analysis restricted only to the presence of MHG.

Conversely, of the societies in the ten regions that did not develop precolonial moralizing gods, only one exceeded the megasociety threshold (the short-lived Inca Empire, social complexity = 0.61). This suggests that, even if moralizing gods do not cause the evolution of complex societies, they may represent a cultural adaptation that is necessary to maintain cooperation in such societies once they have exceeded a certain size, perhaps owing to the need to subject diverse populations in multi-ethnic empires to a common higher-level power[9]. This may explain why moralizing gods spread when large empires conquer smaller—but still complex—societies (for example, the Spanish conquest of the Incas). In some cases, moralizing doctrines may have helped to stabilize empires, while also limiting further expansion; for example, when emperor Ashoka adopted Buddhism and renounced war following his final conquest of the Kalinga Kingdom that established the maximum extent of the Mauryan empire.

Although our results do not support the view that moralizing gods were necessary for the rise of complex societies, they also do not support a leading alternative hypothesis that moralizing gods only emerged as a byproduct of a sudden increase in affluence during a first millennium BC 'Axial Age'[19–22]. Instead, in three of our regions (Egypt, Mesopotamia and Anatolia), moralizing gods appeared before 1500 BC. We propose that the standardization of beliefs and practices via high-frequency repetition and enforcement by religious authorities enabled the unification of large populations for the first time, establishing common identities across states and empires[25,26]. Our data show that doctrinal rituals standardized by routinization (that is, those performed weekly or daily) or institutionalized policing (religions with multiple hierarchical levels) significantly predate moralizing gods, by an average of 1,100 years ($t = 2.8$, d.f. = 11, $P = 0.018$; Fig. 2a). Doctrinal rituals precede moralizing gods in 9 out of the 12 regions analysed, and even precede written records in 6 of these cases (by as much as 4,000 years in the case of Çatalhöyük in Anatolia; see Supplementary Table 2). Although analyses of rates of change of social complexity before and after the appearance of doctrinal rituals do not offer conclusive support for the hypothesis that doctrinal rituals facilitate increasing social complexity (Extended Data Table 3), these data do at least suggest that doctrinal rituals led to the establishment of large-scale religious identities. In the future, higher-quality and higher-resolution archaeological data may allow for a more nuanced understanding of the timing and possible coevolution of the rise of doctrinal rituals and moralizing gods. Such data appear unlikely to affect our primary claim that complex societies preceded moralizing gods, but this is an empirical question open to future testing.

We demonstrate how quantifying cultural characteristics of past societies can contribute to longstanding debates about the evolution of social complexity. Our results suggest that belief in moralizing gods was not the only or even the main factor that enabled the expansion of human societies, but may have occurred along with other features of ritual practices and religion to facilitate cooperation in increasingly complex social systems. In particular, an increase in ritual frequency and doctrinal control may have facilitated the establishment of large-scale collective identities before the spread of beliefs in moralizing gods. Thus, when it comes to the initial rise of social complexity, how you worship may ultimately have been more important than who you worship.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1043-4.

Received: 29 January 2018; Accepted: 19 February 2019;
Published online: 20 March 2019

1. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, London, 1871).
2. Bellah, R. N. *Religion in Human Evolution: From the Paleolithic to the Axial Age* (Harvard Univ. Press, Cambridge, 2011).
3. Purzycki, B. G. et al. Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature* **530**, 327–330 (2016).
4. Watts, J., Sheehan, O., Atkinson, Q. D., Bulbulia, J. & Gray, R. D. Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature* **532**, 228–231 (2016).
5. Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T. & Mace, R. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* **467**, 801–804 (2010).
6. Henrich, J. et al. Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
7. Botero, C. A. et al. The ecology of religious beliefs. *Proc. Natl Acad. Sci. USA* **111**, 16784–16789 (2014).

8. Turchin, P. et al. Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proc. Natl Acad. Sci. USA* **115**, E144–E151 (2018).
9. Norenzayan, A. et al. The cultural evolution of prosocial religions. *Behav. Brain Sci.* **39**, e1 (2016).
10. Norenzayan, A. Big questions about big gods: response and discussion. *Religion Brain Behav.* **5**, 327–342 (2015).
11. Johnson, D. D. P. God's punishment and public goods: a test of the supernatural punishment hypothesis in 186 world cultures. *Hum. Nat.* **16**, 410–446 (2005).
12. Johnson, D. D. P. The wrath of the academics: criticisms, applications, and extensions of the supernatural punishment hypothesis. *Religion Brain Behav.* **8**, 320–350 (2018).
13. Schloss, J. P. & Murray, M. J. Evolutionary accounts of belief in supernatural punishment: a critical review. *Religion Brain Behav.* **1**, 46–99 (2011).
14. Roes, F. L. & Raymond, M. Belief in moralizing gods. *Evol. Hum. Behav.* **24**, 126–135 (2003).
15. Peoples, H. C. & Marlowe, F. W. Subsistence and the evolution of religion. *Hum. Nat.* **23**, 253–269 (2012).
16. Watts, J. et al. Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia. *Proc. R. Soc. B* **282**, 20142556 (2015).
17. Gray, R. D. & Watts, J. Cultural macroevolution matters. *Proc. Natl Acad. Sci. USA* **114**, 7846–7852 (2017).
18. Raffield, B., Price, N. & Collard, M. Religious belief and cooperation: a view from Viking-age Scandinavia. *Religion Brain Behav.* **9**, 2–22 (2019).
19. Baumard, N. & Boyer, P. Explaining moral religions. *Trends Cogn. Sci.* **17**, 272–280 (2013).
20. Baumard, N., Hyafil, A., Morris, I. & Boyer, P. Increased affluence explains the emergence of ascetic wisdoms and moralizing religions. *Curr. Biol.* **25**, 10–15 (2015).
21. Mullins, D. A. et al. A systematic assessment of 'Axial Age' proposals using global comparative historical evidence. *Am. Sociol. Rev.* **83**, 596–626 (2018).
22. Purzycki, B. G. et al. Material security, life history, and moralistic religions: a cross-cultural examination. *PLoS ONE* **13**, e0193856 (2018).
23. Stausberg, M. Big gods in review: introducing Ara Norenzayan and his critics. *Religion* **44**, 592–608 (2014).
24. McKay, R. & Whitehouse, H. Religion and morality. *Psychol. Bull.* **141**, 447–473 (2015).
25. Whitehouse, H. *Modes of Religiosity: A Cognitive Theory of Religious Transmission* (AltaMira, Walnut Creek, 2004).
26. Whitehouse, H., François, P. & Turchin, P. The role of ritual in the evolution of social complexity: five predictions and a drum roll. *Cliodynamics* **6**, 199–216 (2015).
27. Norenzayan, A. & Shariff, A. F. The origin and evolution of religious prosociality. *Science* **322**, 58–62 (2008).
28. Shariff, A. F., Willard, A. K., Andersen, T. & Norenzayan, A. Religious priming: a meta-analysis with a focus on prosociality. *Pers. Soc. Psychol. Rev.* **20**, 27–48 (2016).
29. Turchin, P. et al. Seshat: The Global History Databank. *Cliodynamics* **6**, 77–107 (2015).
30. Murdock, G. P. Ethnographic atlas: a summary. *Ethnology* **6**, 109–236 (1967).

Acknowledgements We thank Q. Atkinson and A. Willard for feedback on an earlier version of the manuscript and E. Postma for discussions on parts of the statistical analyses. We acknowledge the contributions of our team of research assistants, post-doctoral researchers, consultants and experts. See http://www.seshatdatabank.info for a comprehensive list of private donors, partners, experts and consultants. This work was supported by an ESRC Large Grant entitled 'Ritual, Community, and Conflict' (REF RES-060-25-0085), a John Templeton Foundation grant to the Evolution Institute entitled 'Axial-Age Religions and the Z-Curve of Human Egalitarianism', a Tricoastal Foundation grant to the Evolution Institute entitled 'The Deep Roots of the Modern World: The Cultural Evolution of Economic Growth and Political Stability', an Advanced Grant ('Ritual Modes: Divergent modes of ritual, social cohesion, prosociality, and conflict', grant agreement no. 694986) and a Starter Grant ('The Cultural Evolution & Ecology of Institutions', grant agreement no. 716212) from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme, an award from the Templeton World Charity Foundation entitled 'Cognitive and Cultural Foundations of Religion and Morality' (TWCF0164), a Keio Research Institute at SFC Startup Grant, a Keio Gijuku Academic Development Fund Individual Grant and a grant from the European Union Horizon 2020 Research and Innovation Programme (grant agreement no. 644055 (ALIGNED, www.aligned-project.eu)).

Reviewer information *Nature* thanks Carol Ember, Peter J. Richerson and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions H.W., P.F. and P.E.S. designed the study, with input from P.T., T.E.C., K.C.F. and R.M.R.; E.C. and R.P. coded the religion and ritual data, with additional input from J.L., J.B., B.t.H and A.C.; P.E.S. analysed the data, with input from H.W., P.F., P.T., T.E.C. and R.M.R.; P.E.S., H.W. and P.F. drafted the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1043-4.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1043-4.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to P.E.S.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Seshat: Global History Databank overview. Seshat (http://seshatdatabank.info/) is a vast database of information about global history from the end of the Palaeolithic period up to the Industrial Revolution. Eventually, it is intended that Seshat will cover the history of all past human societies; however, initially the goal has been to capture as much diversity in global history as possible. We therefore created a stratified sample of past societies by identifying ten world regions distributed as widely as possible across the Earth's surface and within each of those regions designated three 'natural geographical areas' (NGAs) with discrete ecological boundaries, on average about 10,000 km^2 in size, thus creating an initial sampling scheme of 30 such areas around the world. To maximize diversity in the sample, for each world region we chose one NGA in which social complexity emerged early (for example, Egypt), one in which it arose relatively recently (for example, Iceland) and one in which it emerged somewhere in the middle of the range (for example, Japan)[29]. The 30 regions and their selection rationale have previously been published[29], before the start of data collection. Our aim was to maximize variability in our global sample while minimizing historical relationships between cultures.

Data on political systems (polities) that emerged and persisted in each of the NGAs were then gathered and entered into Seshat in a continuous time series at 100-year intervals, going back as far into the history of that area as scholarly literature would allow (up to a maximum of roughly 10,000 years before present). In the case of NGAs that contained clusters of very small-scale polities that share a similar culture but are not under a single system of jurisdictional control, we refer to these as 'quasi-polities' and code information on all of them generically, unless information is available that would allow us to differentiate between these polities.

All variables for which data have been gathered and entered into Seshat are derived from a Seshat Codebook that can be accessed and downloaded (http://seshatdatabank.info/methods/codebook). The Codebook was designed by, and is continually updated and extended in consultation with, a large network of professional historians, archaeologists, anthropologists and other specialists whom we refer to as 'Seshat experts'. Most variables in Seshat require the data to take the form of a number or numerical range or they specify a feature that can be coded as absent, present or unknown (additionally coding items as 'inferred present' or 'inferred absent', where the evidence permits). The first step in data entry was for trained research assistants to gather and input easily acquired data, and at the same time to compile lists of data that are more difficult to interpret and that require input from Seshat experts. Especially during the early phases of data entry, variables in the codebook were revised and improved through continuous discussions between research assistants and Seshat experts. All data are linked to scholarly sources, including peer-reviewed publications and personal communications from established authorities. On occasions when Seshat experts disagree on a particular coding, we kept a record of disagreements so that analyses could be run taking into account contrasting interpretations. Once used for the purposes of data analysis and publication, that version of the dataset was 'frozen' so that it could be inspected by others and used for the purposes of replication. Nevertheless, the data in Seshat continually evolves, as new sources are discovered and as new Seshat experts contribute additional layers of interpretation.

The data analysed in the present paper focused on those sectors of the Seshat Codebook concerned with social complexity, religion and ritual. A full account of the social complexity variables has previously been published[8], using 51 variables (Extended Data Table 5) associated with population size, hierarchy, territory, governance, bureaucracy, infrastructure, record keeping, economic development and other domains that were previously identified as potentially relevant measures of social complexity. This required engagement with a wide-ranging body of literature on social complexity. Because previous researchers disagreed about which dimensions of social complexity were the most important to emphasize (for example, number of jurisdictional levels versus more-horizontal forms of complexity; autocracy versus democracy; diversity of specialist roles versus centralized coordination), we included proxies for all potentially relevant measures of social complexity that had been identified in the literature. This inclusive strategy was designed to allow us to investigate whether these different characteristics exhibited strong relationships with each other and whether a single principal component captured most of the observed variation. Our analyses confirmed that both are indeed the case. Furthermore, we found that different characteristics of social complexity were highly predictable across different world regions[8].

Whereas previous research has proposed an association between the rise of moralizing gods and the evolution of social complexity, measures used in the past to capture the latter have been comparatively crude. Variable selection and inclusion for moralizing gods was informed by existing literature on so-called big gods, MHG and BSP, as well as psychological and cross-cultural comparative research on the hypothesized link between belief in moralizing gods and large-scale cooperation.

Data collection for the religion and ritual variables involved matching each fully trained research assistant with one or more Seshat experts. Seshat experts provided guidance on how to delineate the temporal and geographical boundaries of the polity, assembled an initial reading list and—where necessary—helped to interpret some of the key historiographical debates associated with the variables. Research assistants then populated the variables with data and presented this to the Seshat experts for review. The comments and suggestions made by the experts were then implemented by the research assistants. The next stage required a second team of fully trained research assistants to go over the gathered data and to conduct a series of quality checks, including vetting of the footnotes and the use of correct syntax for the machine-readable part of the data. Finally, this checked dataset was given to the Seshat experts for review. The coding of religion and ritual data required the input of experts every step of the way, given the frequent need for complex and nuanced interpretation of the evidence. By contrast, the data required for the social complexity variables frequently consisted of facts that research assistants could procure with less supervision, allowing expert input and review to occur at a later stage of the process.

Data coding. *Social complexity*. The 51 variables used to construct the overall social complexity measure are shown in Extended Data Table 5. These variables were chosen because they reflect common features associated with social complexity and were grouped into nine complexity characteristics (polity population size, capital population size, polity territory size, hierarchy, infrastructure, government, information systems, texts and money). Details of coding definitions for these variables have previously been published[8,29].

MHG. For consistency with previous studies that have generally used the MHG variable from the Ethnographic Atlas[30], the presence of MHG was coded as a binary variable on the basis of the original definition of a high god as "a spiritual being who is believed to have created all reality and/or to be its ultimate governor, even though his sole act was to create other spirits who, in turn, created or control the natural world"[31]. The following categories were used: (1) absent or not reported, (2) present but not active in human affairs, (3) present and active in human affairs but not supportive of human morality and (4) present, active and specifically supportive of human morality[11,32].Thus, a coding of high gods as present, active and specifically supportive of human morality was coded as a MHG being present, whereas all other types were coded as absent.

BSP. The terms big gods and MHG are sometimes used interchangeably[17], but can have different connotations. The term MHG was previously developed and defined, and it was proposed that high gods were associated with social complexity (regardless of their moral concern)[31]. This definition of the MHG variable was incorporated into the Ethnographic Atlas[30], resulting in it being widely used in cross-cultural research. These ideas were subsequently extended in the supernatural punishment hypothesis[11,12,33], in which the focus lay on the mechanism of morality enforcement rather than high gods; however, the MHG variable was used for testing because of the availability of previous research using this definition. The ideas have been further developed[9,10,27,34] to include various additional mechanisms, most notably including cultural group selection to explain the rapid spread of moralizing gods without accompanying genetic changes.

The term big gods (defined as "powerful, omniscient, interventionist, morally concerned gods"[34]) was originally the title of a monograph describing this theory. Later, however, this definition was relaxed and it was emphasized that the term big gods was a rhetorical device intended to include a broad range of morally concerned supernatural agents, not only MHGs: "…powerful, all-knowing and morally concerned supernatural agents who are believed to monitor social interactions and to reward and sanction behaviours in ways that contribute to the cultural success of the group, including practices that effectively transmit the faith. Rhetorically, we call these 'Big Gods'; however, we alert readers that we are referring to a multidimensional continuum of supernatural agents in which big gods occupy a particular corner of the space"[9].

Subsequently an additional variable was developed that was called broad supernatural punishment[16], which arguably more closely matches this relaxed definition of big gods than does the traditional MHG variable. BSP was defined as follows: "For BSP to be coded as present in a culture there must be the concept of a supernatural agent or process that reliably monitors and punishes selfish actions, and this concept must (i) be widely advocated within the community, (ii) involve punishment of a broad range of selfish behaviours and (iii) apply to a wide range of community members."[16].

Because selfish actions can occur in a variety of domains, Seshat subdivides the types of supernatural enforcement of morality based on nine proposed categories of morality[35,36]. For this study, we focused on three domains that are relevant to the establishment of large-scale cooperation: (1) fairness (sharing of resources, such as dividing disputed resources, bargaining or redistribution of wealth); (2) reciprocity (for example, fulfilling contracts, returning gifts, repaying debts or upholding trust); and (3) in-group loyalty (the need to remain loyal to unrelated members of the same group; for example, helping coreligionists or going to war for one's group).

BSP was coded as present if at least one of these three sub-types of selfish actions was supernaturally enforced.

Our robustness analyses, which evaluated BSP and MHG separately (Extended Data Table 4) suggest that—rather than moralizing gods following a general pattern of evolution from small (BSP) to big (MHG)—the presence or absence of high gods independent of their moralizing status has little functional relationship with social complexity, and instead appears largely contingent on history and geography. In regions such as southern and eastern Asia, BSP in the form of karmic religions (Buddhism and Hinduism) remains the dominant form of moralizing gods, whereas in regions such as Europe and Africa moralizing Abrahamic MHGs were commonly adopted or imposed without any intermediate evolution through a BSP stage.

Doctrinal rituals. The modes of religiosity hypothesis focuses on two factors that facilitate standardization of a body of beliefs and practices. First, high frequency (for example, daily or weekly) collective rituals facilitate easy detection of deviations from the orthodox canon. Second, religious hierarchy enables enforcement of authorized belief and practice. Seshat codes five different types of rituals: the most frequent, most widespread, largest scale, most euphoric and most dysphoric rituals. For each ritual, frequency is coded as daily, weekly, monthly, seasonally, yearly, generationally or once-in-a-lifetime. Seshat also encodes levels of religious hierarchy. One represents no levels of religious hierarchy beyond the local priest or shaman, whereas higher numbers represent multiple levels of hierarchy (for example, senior priests or high druids).

Making inferences about prehistoric rituals requires using various measurable archaeological proxies. Previous research has established that both frequent rituals and multi-level religious hierarchies tend to co-occur with other features of doctrinal rituals (for example, low arousal)[25,37–39]. Not all of these features can always be found in the archaeological record, so in this paper we use the appearance of either religious hierarchy or frequent rituals as proxies for the appearance of doctrinal rituals. Doctrinal rituals were thus coded as present if the most frequent ritual occurred weekly or daily, or if there was evidence of multiple levels of religious hierarchy.

Separate re-analyses were also conducted in which doctrinal rituals were defined based only on ritual frequency and only on religious hierarchy. In both cases, doctrinal rituals still preceded moralizing gods by an average of over 200 years, although this difference only remained significant when using religious hierarchy as a proxy for doctrinal ritual practices (religious hierarchy: mean = 991 years, $t = 2.4$, d.f. = 11, $P = 0.035$; ritual frequency: mean = 210 years, $t = 1.1$, d.f. = 11, $P = 0.30$).

Note that we coded only aspects of ritual practices and religion associated with the official cult, and so the rituals of interest were not necessarily polity-wide but could be largely or wholly restricted to elite groups.

Data collation. The process of data collection for the MHG, BSP and doctrinal rituals variables involved matching each fully trained research assistant with one or more experts (recognized authorities on the polity in question, typically holding a relevant doctorate and occupying a faculty position in a university). Experts provided guidance on how to delineate the temporal and geographical boundaries of the polity, assembled an initial reading list and—where necessary—helped to interpret some of the key historiographical debates associated with the variables. Research assistants then populated the variables with data and presented these data to the experts for review. The comments and suggestions made by the experts were then implemented by the research assistants. The next stage required a second team of fully trained research assistants to conduct a series of quality checks, including vetting of the footnotes (which currently reference over 2,000 unique sources) and the use of correct syntax for the machine-readable part of the data. Finally, this checked dataset was offered to the experts for review. By contrast, the data required for the social complexity variables frequently consisted of facts that research assistants could procure with less supervision, allowing expert input to occur at a later stage of the process. Data vetting in Seshat is a continuously ongoing dynamic process that includes incorporation of disagreement among experts within the project and input from external experts via our open-access interface.

There is room for reasonable disagreement about the most effective way of gathering data about world history, particularly regarding the role of expert contributors[40,41]. An alternative approach would be to have every single data point signed off by a single recognized expert, perhaps even without requiring further citations. We trialed such an expert-driven approach to data entry during initial phases of our project but found it took too long to source experts and have them enter the data required. Instead we found that faster progress could be made using the approach described above and having multiple points at which the data were examined and vetted. To this end, we have made not only all our data but also all of the metadata and references supporting these data available for everyone to examine and comment on. Rather than relying on the authority of a single expert for each entry, Seshat involves regional experts to help to guide data collection and assess the quality of our data and metadata as one of several complementary components in our quality control approach, which also includes incorporating disagreement among multiple experts.

Analyses. To ensure consistency and comparability in our analyses, we sampled polities at 100-year intervals, sampling whichever polity happened to occupy a given region at AD 100, AD 200, AD 300 and so on, and not including polities that existed only between century boundaries[8] (see Supplementary Information for details and examples regarding the temporal sampling procedure). All analyses were performed in R v.3.4.1[42]. All *P* values reported are from two-tailed analyses.

Quantifying social complexity. To create an overall measure of social complexity, we took a previously published approach based on principal component analysis (PCA)[8] and applied it to the latest available data from Seshat. This method aggregates the 51 social complexity variables (Extended Data Table 5) into nine complexity characteristics and then analyses them using PCA.

PCA is a commonly used tool for dimension reduction—in this case we have nine different aggregated variables that we want to reduce to a single variable that best captures social complexity. However, we obtain the same conclusions even without using PCA, regardless of which of the nine complexity characteristics we choose as a proxy for social complexity (Extended Data Table 4).

As previously shown, these different complexity characteristics turn out to be highly correlated and all load heavily onto a single principal component that captures 76% of the variance in the individual complexity characteristic variables. Our approach uses multiple imputation[43] to account for missing data, uncertainty and expert disagreement by imputing data based on a range of possible values and averaging the results over the course of 20 imputations. The results of this approach have proven highly robust to a number of different modelling assumptions[8] (Supplementary Information). Full details of this approach and justifications for selecting the social complexity variables can be found in a previously published paper[8]. We previously carried out a number of robustness checks[8], including cross-validation analysis and bootstrap resampling to assess whether our PCA methods were robust to spatio-temporal autocorrelation. Specifically, *k*-fold cross-validation showed that our multiple imputation methods accurately predicted complexity characteristic values when each geographical region was systematically removed from the analysis, and bootstrapping showed that removing different geographical regions and time periods did not affect our PCA results[8] (see Supplementary Information for full details).

Extensions of PCA (for example, generalized low-rank models[44], spatio-temporal PCA[45] and singular spectrum analysis[46]) may be worth considering in future analyses as alternative methods of accommodating binary variables and spatio-temporal autocorrelation. Note, however, that the subsequent regression analyses performed in this paper explicitly control for spatial, temporal and phylogenetic autocorrelation. More importantly, our current results consistently failed to support the temporal sequence of the moralizing gods hypothesis across all geographical regions (Extended Data Table 1 and Extended Data Fig. 1), and robustness analyses using each of the nine complexity characteristics independently without performing PCA also confirmed our main findings (Extended Data Table 4). This confirms that our primary finding that complex societies precede moralizing gods cannot be an artefact of autocorrelation in our PCA methods.

Logistic regression. To examine the association between moralizing gods and social complexity while controlling for non-independence in our data owing to spatio-temporal autocorrelation and historical connections between cultures[47], we fitted a logistic regression model to the data. A detailed description of this model has previously been published along with extensive validation of its robustness when applied to Seshat data[48]. This approach stems from the field of nonlinear dynamical systems, and is similar to the concept of Granger causality[49,50] (which is commonly used in economics), in that both use linear models with time-lagged variables.

Our approach is similar to a previously published study[7], except that we use more fine-grained measures of geographical diffusion and linguistic similarity, and also incorporate temporal information, as follows:

$$Y_{i,t} = a + \sum_{\tau} b_{\tau} Y_{i,t-\tau} + c \sum_{i \neq j} \exp\left[-\frac{\delta_{i,j}}{d}\right] Y_{j,t-1} + h \sum_{i \neq j} w_{i,j} Y_{j,t-1} + \sum_{k} g_k X_{k,i,t-1} + \epsilon_{i,t}$$

in which $Y_{i,t}$ is the binary variable encoding the presence or absence of moralizing gods in location i at time t. The time step $\Delta t = 100$ years. Starting from the first term on the right-hand side, a is the regression constant (intercept). The next term captures the influences of past history (autoregressive terms), with $\tau = 1, 2, \ldots$ indexing time-lagged values of Y (as time is measured in centuries, $Y_{i,t-1}$ refers to the presence or absence of moralizing gods 100 years before t). The third term represents potential effects resulting from geographical diffusion[51,52]. We use a negative-exponential form to relate the distance between society i and society j ($\delta_{i,j}$) to the influence of j on i because—in contrast to a linear kernel—a negative exponential does not become negative at very long $\delta_{i,j}$; it instead approaches

0 smoothly. We avoid the problem of endogeneity by using time-lagged $Y_{j,t-1}$. Thus, the third term is a weighted average of the occurrence of moralizing gods in the vicinity of society i at the previous time step, with weights falling off to 0 as distance from i increases. Parameter d measures how steeply the influence falls with distance, and was set to $d = 1{,}000$ km after optimizing the Akaike information criterion value using 200-km increments from 200 to 2,000 km ($d = 200, 400, 600, \ldots, 2{,}000$ km). Parameter c is a regression coefficient measuring the importance of geographical diffusion. Detecting autocorrelations owing to shared cultural history (next term) is done analogously, except w now represents the weight due to linguistic similarity (set to 1 if societies i and j share the same language, 0.5 if they are in the same linguistic genus, 0.25 if they are in the same linguistic family, and 0 if they are in different linguistic families; linguistic genera and families were taken from Glottolog[53] and the World Atlas of Language Structures[54]). The rest of the right-hand side represents effects of predictor variables $X_{k,i,t-1}$ (time-lagged); g_k are regression coefficients and $\varepsilon_{i,t}$ is the error term. This approach allows us to investigate the effects of the predictor variable (social complexity, calculated above via PCA), while controlling for serial autocorrelations, spatial diffusion and autocorrelations due to the shared cultural history. The regression results are detailed in Extended Data Table 2.

Comparison of before and after moralizing gods. To more directly examine the direction of causality predicted by the moralizing gods hypothesis, we created a time series of social complexity over time for all 12 regions for which social complexity data were available both before and after the appearance of moralizing gods (Extended Data Fig. 1). We then compared rates of change in social complexity over time before and after moralizing gods using sliding time windows. First, we compared rates of change using a 100-year window (that is, comparing the rate for the 100 years before the appearance of moralizing gods with the rate for the 100 years after), then repeated this using a 200-year window, 300-year window and so on, up to a maximum of between 700- and 3,900-year windows depending on the region (different regions have different time depths of data available for making these comparisons).

The region with the shallowest time depth was in Mali (±700 years before/after the appearance of Islam around 1100 AD (400 AD–1800 AD)), whereas the region with the deepest time depth was in Iran (±3,900 years before/after evidence of the moralizing Mesopotamian sun god Shamash around 2200 BC (6100 BC–1700 AD)). If we used all available data (up to ±3,900 years), we risked weighting the analyses too heavily towards regions such as Iran with deep time depths, whereas using only a consistent upper limit of a maximum of ±700 years risks throwing away too much data. As a compromise, we conducted analyses using an intermediate upper limit of a maximum of ±2,000 years (Fig. 2b), but also repeated the analyses using extreme upper limits of ±700 years and ±3,900 years (see 'Robustness analyses'). All of these choices produced qualitatively identical results (Extended Data Table 4).

Note that these analyses do not attempt to construct a single average rate of change before moralizing gods, a single average rate of change after moralizing gods, and compare these average rates. We cannot assume such a constant rate of change—and, indeed, Fig. 2a clearly shows that rates of change are not constant. Instead, these analyses calculate a difference value for each time window (for example, subtracting the rate of change for the 100-year period before moralizing gods from the rate of change for the 100-year period after, and then doing the same for a ±200-year period, and so on). The key prediction of the moralizing gods hypothesis is that these difference values should tend to be positive (that is, for a given time window, the rate of change after moralizing gods should be greater than the rate before). However, Fig. 2b demonstrates that—in fact—the distribution of difference values was significantly negative (paired t-tests, $P < 10^{-9}$).

Robustness analyses. To explore the robustness of our results to modelling assumptions, we ran the following robustness analyses.

To ensure that the analyses are not affected by the fact that religious hierarchy is included as one of the social complexity variables in addition to being one of the variables used to define doctrinal mode, we reran the analyses after removing the religious hierarchy variable from the social complexity variables. We chose to do this for robustness analyses rather than the primary analysis in order to use the same 51 social complexity variables used in our previously published studies[8,48] for consistency.

To ensure that the observed plateauing of social complexity was not simply an artefact of a ceiling effect wherein polities 'max out' certain variables, we reran the analyses twice after splitting the social complexity variables in two subsets. The scale subset contained only the subset of seven social complexity variables for which there was no theoretical maximum value (from the categories polity population, polity territory, capital population and hierarchy). The non-scale subset contained the remaining 44 social complexity variables for which there was a theoretical maximum that could be attained once all our variables were present in a society (from the categories government, money, infrastructure, information systems and texts).

To examine whether our results were affected by the definition of moralizing gods, we reran the analyses limiting the definition of moralizing gods exclusively to MHG, rather than the more inclusive definition of BSP used in the primary analysis.

Our primary analysis treated moralizing gods as being present from the beginning of the polity in which they appeared. To ensure that our analyses were not affected by dating uncertainty, we reran the analyses randomly resampling to treat moralizing gods as appearing at some point from within the full date range of this polity (for example, 2900–2700 BC for Egypt).

Our primary analysis used time windows of up to 2,000 years before and after the appearance of moralizing gods, because 2,000 years was intermediate between the maximum time window for the region with the shallowest time depth (±700 years for Mali) and the deepest time depth (±3,900 years for Iran). To examine whether our results were affected by the depth of the time window used, we reran analyses using consistent time windows for each region of up to 700 years before/after moralizing gods (because ± 700 was the maximum time window possible for Mali), and also using the full time window available for each region (that is, as wide as 3,900 years for Iran).

To ensure that our results were not affected by possible autocorrelation in our use of PCA to extract a measurement of social complexity, we reran the analysis nine times using each of the nine individual complexity characteristics as a measure of social complexity without performing any PCA.

All of these robustness analyses (16 in total) produced qualitatively identical results in which the rate of increase of social complexity was significantly greater before the appearance of moralizing gods than afterwards (more than double in all cases; Extended Data Table 4), which confirms that our primary conclusion that complex societies precede moralizing gods is highly robust.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

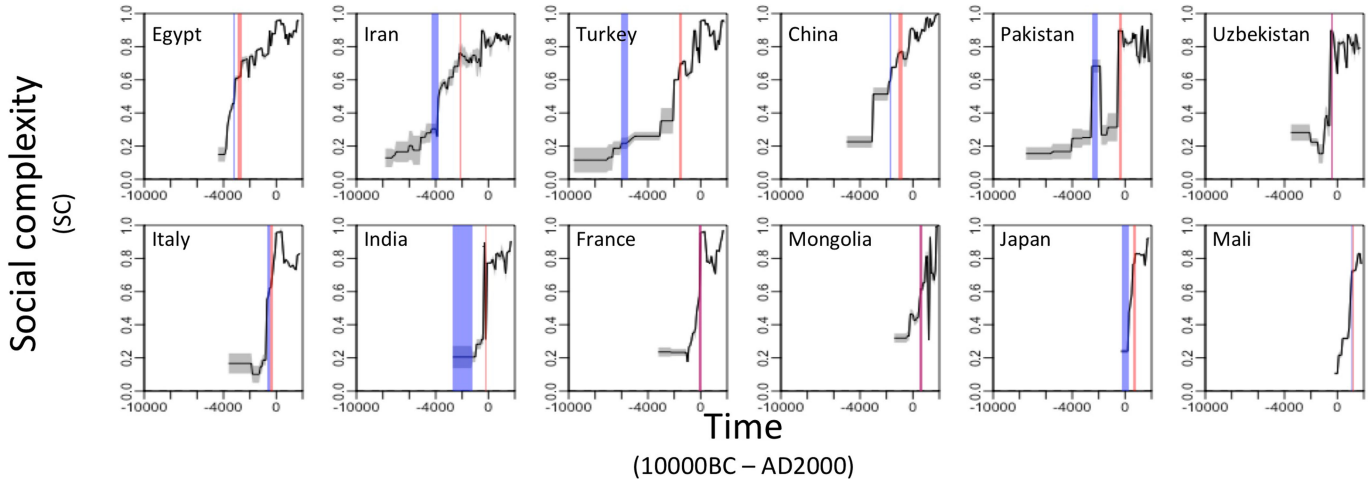
Code availability. Source code is available online at http://github.com/pesavage/moralizing-gods.

Data availability

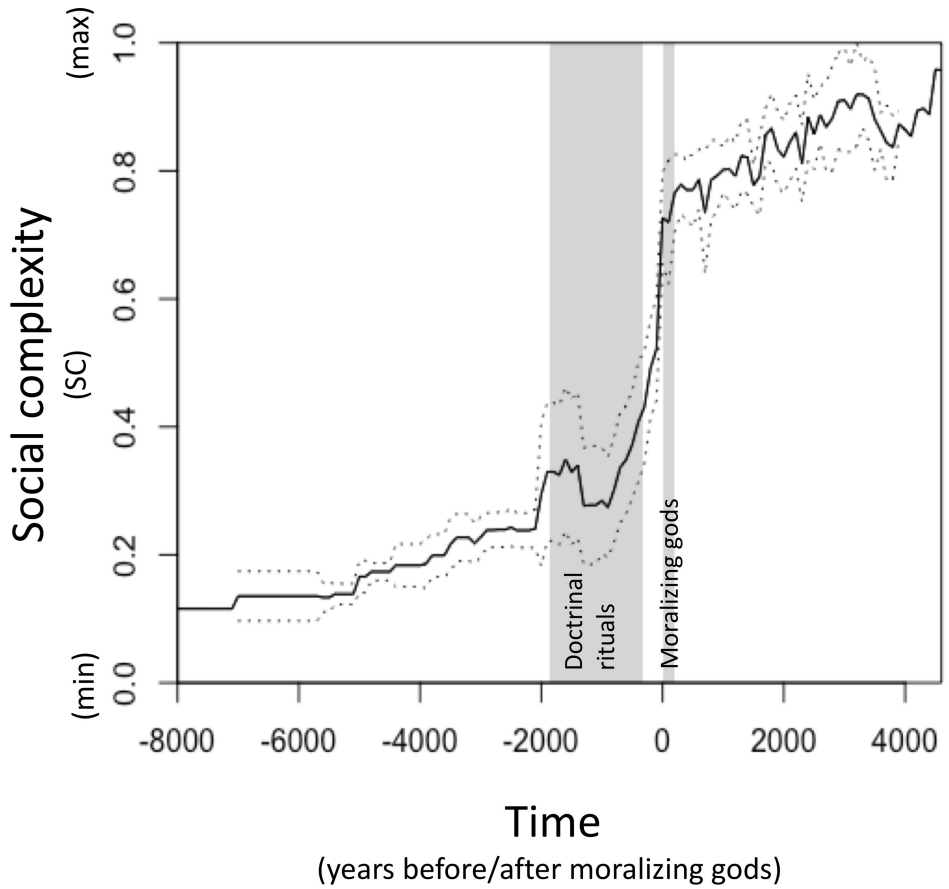
The full machine-readable dataset is available as Supplementary Data 1, and at http://seshatdatabank.info/datasets. Full coding data with detailed explanations and references are available at http://seshatdatabank.info/data, and are summarized in Supplementary Table 2. The data include the coded levels of uncertainty and disagreement, the textual explanations and the references for each of the variables for all polities used in our analysis. These webpages also make it possible to comment on each of our data points and suggest additions or corrections and thus provide an up-to-date and dynamic dataset that undergoes continual improvement by members of the Seshat team and external scholars. To maximize transparency, we have tied each cluster of variables to the names of the research assistants who gathered the data, and to the names of the experts who reviewed the data.

31. Swanson, G. E. *The Birth of the Gods: The Origin of Primitive Beliefs* (Michigan Univ. Press, Ann Arbor, 1960).
32. Divale, W. *Pre-Coded Variables for the Standard Cross-Cultural Sample* Vols 1 and 2 (York College, New York, 2000).
33. Johnson, D. *God is Watching You: How the Fear of God Makes Us Human* (Oxford Univ. Press, New York, 2016).
34. Norenzayan, A. *Big Gods: How Religion Transformed Cooperation and Conflict* (Princeton Univ. Press, Princeton, 2013).
35. Haidt, J. & Graham, J. When morality opposes justice: conservatives have moral intuitions that liberals may not recognize. *Soc. Justice Res.* **20**, 98–116 (2007).
36. Curry, O. S. in *The Evolution of Morality* (eds Shackelford, T. K. & Hansen, R. D.) 27–51 (Springer, Cham, 2016).
37. Whitehouse, H. & Lanman, J. The ties that bind us: ritual, fusion, and identification. *Curr. Anthropol.* **55**, 674–695 (2014).
38. Whitehouse, H. & Martin, L. H. *Theorizing Religions Past: Archaeology, History, and Cognition* (AltaMira, Blue Ridge Summit, 2004).
39. Atkinson, Q. D. & Whitehouse, H. The cultural morphospace of ritual form: examining modes of religiosity cross-culturally. *Evol. Hum. Behav.* **32**, 50–62 (2011).
40. Slingerland, E. & Sullivan, B. Durkheim with data: the Database of Religious History. *J. Am. Acad. Relig.* **85**, 312–347 (2017).
41. Wade, L. Birth of the moralizing gods. *Science* **349**, 918–922 (2015).
42. R Core Team. R: A Language and Environment for Statistical Computing. https://www.r-project.org/ (2015).
43. Rubin, D. B. *Multiple Imputation for Nonresponse in Surveys* (Wiley, New York, 1987).
44. Udell, M., Horn, C., Zadeh, R. & Boyd, S. Generalized low rank models. *Found. Trends Mach. Learn.* **9**, 1–118 (2016).
45. Stahlschmidt, S., Härdle, W. K. & Thome, H. An application of principal component analysis on multivariate time-stationary spatio-temporal data. *Spat. Econ. Anal.* **10**, 160–180 (2015).
46. Hassani, H. Singular spectrum analysis: methodology and comparison. *J. Data Sci.* **5**, 239–257 (2007).

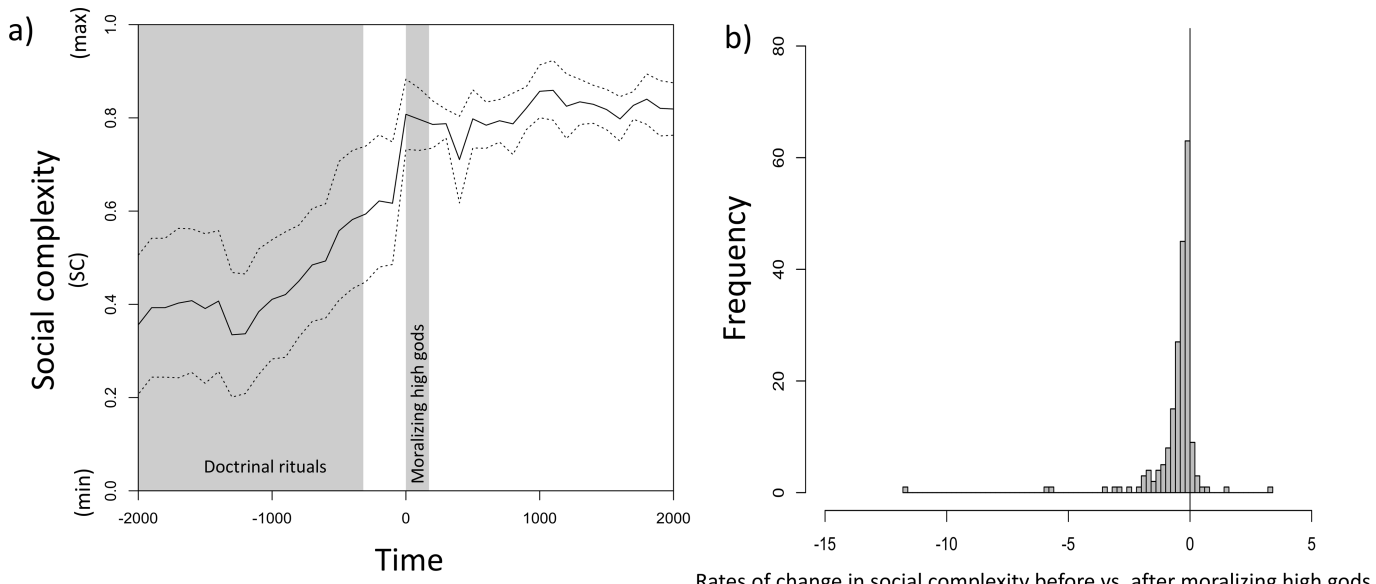
47. Mace, R. & Pagel, M. The comparative method in anthropology. *Curr. Anthropol.* **35**, 549–564 (1994).
48. Turchin, P. Fitting dynamical regression models to Seshat data. *Cliodynamics* **9**, 25–58 (2018).
49. Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438 (1969).
50. Kleinberg, S. *Causality, Probability, and Time* (Cambridge Univ. Press, Cambridge, 2013).
51. Eff, E. A. & Dow, M. M. How to deal with missing data and Galton's problem in cross-cultural survey research: a primer for R. *Struct. Dyn.* **3**, 1–29 (2009).
52. Eff, E. A. & Routon, P. W. Farming and fighting: an empirical analysis of the ecological-evolutionary theory of the incidence of warfare. *Struct. Dyn.* **5**, 1–33 (2012).
53. Hammarström, H., Forkel, R., Haspelmath, M. & Nordhoff, S. Glottolog version 2.3 http://glottolog.org (2014).
54. Dryer, M. S. & Haspelmath, M. *The World Atlas of Language Structures online* http://wals.info (Max Planck Institute for Evolutionary Anthropology, 2013).



Extended Data Fig. 1 | Social complexity time series for individual regions. The 12 regions for which social complexity data are available both before and after the appearance of moralizing gods are shown. Vertical bands represent the period in which the first evidence of moralizing gods (red) and doctrinal rituals (blue) appeared. Grey shading represents 95% confidence intervals based on a PCA using multiple imputation[8].



Extended Data Fig. 2 | Full time series showing mean social complexity over time before and after the appearance of moralizing gods. $n = 12$ regions with data before and after the appearance of moralizing gods. Social complexity has been scaled so that the society with the highest social complexity (Qing Dynasty, China, around AD 1900) has a value of 1 and the society with the lowest social complexity (Early Woodland, Illinois, USA, around 400 BC) has a value of 0. Vertical bands represent the period in which moralizing gods and doctrinal rituals first appeared. All errors represent 95% confidence intervals, with the exception of the vertical bar for moralizing gods, which represents the mean duration of the polity in which moralizing gods appeared (because times are normalized to the time of first evidence of moralizing gods, and there is thus no variance in this parameter). Lack of confidence intervals indicates data from only a single region. This figure is identical to Fig. 2a, except that it also includes all available data before and after moralizing gods, rather than being restricted to a window of 2,000 years before and after.



Extended Data Fig. 3 | Social complexity before and after the appearance of MHG. This is a version of Fig. 2 in which analyses are restricted to only MHG, rather than the broader definition of moralizing gods used in Fig. 2 and elsewhere (which includes BSP as well as MHG). **a**, Time series showing the mean social complexity over time for 2,000 years before and after the appearance of MHG. $n = 10$ regions with social complexity data for before and after moralizing high gods. Social complexity has been scaled so that the society with the highest social complexity (Qing Dynasty, China, around AD 1900) has a value of 1 and the society with the lowest social complexity (Early Woodland, Illinois, USA, around 400 BC) has a value of 0. Vertical bands represent the period in which MHG and doctrinal rituals first appeared. All errors represent 95% confidence intervals, with the exception of the vertical bar for MHG, which represents the mean duration of the polity in which MHG appeared (because times are normalized to the time of first evidence of MHG and there is therefore no variance in this parameter). **b**, Histogram of differences in rates of change in social complexity after minus before the appearance of MHG ($n = 158$ time windows from the 10 regions). The *y* axis represents the number of time windows out of 158.

Extended Data Table 1 | Timing and rates of change in social complexity before and after the earliest precolonial evidence of moralizing gods

Modern country name	Religious system	Polity name/dates	Moralizing gods type	Rate of change in social complexity before vs. after moralizing gods Paired t	n
Egypt	Pharaonic religion	2nd Dynasty (2900-2700BCE)	MHG	-5.1***	15
Iran	Akkadian religion	Dynasty of Akkad (2250-2090BCE)	BSP	-8.4***	20
Turkey	Hittite religion	Old Kingdom of Hatti (1650-1500BCE)	BSP	-6.4***	20
China (Henan)	Zhou religion	Western Zhou Dynasty (1040-770BCE)	MHG	-7.3***	20
Pakistan	Zoroastrianism	Achaemenid Empire (550-330BCE)	MHG	-3.6**	20
Uzbekistan	Zoroastrianism	Achaemenid Empire (520-330BCE)	MHG	-4.3***	20
Italy	Roman religion	Early Roman Republic (510BCE-270BCE)	BSP	-3.2**	20
India (Karnataka)	Buddhism	Mauryan Empire (300-200BCE)	BSP	-3.7***	18
France	Celtic religion	La Tène Gaul (170-30BCE)	BSP	1.2	17
Cambodia	Buddhism	Early Funan (220-540CE)	BSP	NA	0
Mongolia	Buddhism	Rouran Khaganate (300-550CE)	BSP	0.2	14
Yemen	Judaism	Himyarite Kingdom (380-530CE)	MHG	NA	0
Japan	Buddhism	Asuka period (540-710CE)	BSP	-7.1***	9
Indonesia (Java)	Buddhism	Medang Kingdom (730-1020CE)	BSP	NA	0
Iceland	Norse religion	Icelandic Commonwealth (930-1260CE)	BSP	NA	0
Mali	Islam	Kingdom of Ghana (1100-1200CE)	MHG	-6.1***	7
Ghana	Akan religion	Pre-Ashanti period (1500-1700CE)	MHG	NA	0
Indonesia (Kalimantan)	Iban religion	Pre-Brooke Raj period (1650-1840CE)	BSP	NA	0
USA (Hawaii)	Hawaiian religion	Pre-US period (1780-1820CE)	BSP	NA	0
Micronesia (Chuuk)	Chuuk religion	Pre-German period (1770-1890CE)	BSP	NA	0
Russia (Yakutsk)	Sakha religion	Pre-Russian period (1400-1640CE)	None	NA	0
India (Garo)	Garo religion	Pre-British period (1770-1870CE)	None	NA	0
China (Yunnan)	Hmong religion	Late Qing period (1700-1900CE)	None	NA	0
Papua New Guinea (Oro)	Orokaiva religion	Pre-British period (1730-1880CE)	None	NA	0
USA (Finger Lakes)	Iroquois religion	Pre-British period (1570-1710CE)	None	NA	0
USA (Missouri)	Illinois religion	Illinois Confederacy (1640-1720CE)	None	NA	0

For locations without precolonial concepts of moralizing gods, the polity represents the latest polity analysed. See Supplementary Table 2 and http://seshatdatabank.info/data for details and references. Rates of change before and after moralizing gods were compared using paired *t*-tests on up to 20 time windows (100–2,000 years before and after the appearance of moralizing gods) for all 12 regions with social complexity data available both before and after the appearance of moralizing gods. Negative *t*-values represent higher rates of change before moralizing gods. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Extended Data Table 2 | Logistic regression results predicting moralizing gods

Parameter	Coefficient estimate (± SE)	z-value	*P* [Pr(>|z|)]
(Intercept)	-7.9 ± 0.9	-8.7	<.001
Social complexity (SC)	9.9 ± 1.5	6.8	<.001
Time (100-year lag)	3.7 ± 0.7	5.5	<.001
Language phylogeny	11.6 ± 5.5	2.1	.04
Geographical proximity	-2.5 ± 1.3	-2.0	.04
Time (200-year lag)	0.9 ± 0.7	1.3	.19

The model includes parameters for social complexity and for geographical, temporal and cultural relationships, ordered by absolute *z*-value (see Methods for details). SE, standard error. Pr, probability.

Extended Data Table 3 | Analyses with doctrinal rituals instead of moralizing gods as the dependent variable

	Logistic regression			Change in SC pre-/post- doctrinal rituals			
	z	***P***	**n (centuries)**	**Mean pre:post ratio**	**Paired t**	***P***	**n (century time-windows)**
1. Doctrinal ritual defined via religious hierarchy or ritual frequency	7.4	<.001	801	2.0	-4.6	<.001	189
2. Doctrinal ritual defined via religious hierarchy only	7.9	<.001	801	2.4	-5.2	<.001	203
3. Doctrinal ritual defined via ritual frequency only	6.2	<.001	801	7.7	-11.2	<.001	202

See Supplementary Tables 3–5 for full regression results.

Extended Data Table 4 | Robustness analyses modifying modelling assumptions of the analyses

	Logistic regression			Change in SC pre-/post- doctrinal rituals			
	z	*P*	n (centuries)	Mean pre:post ratio	Paired t	*P*	n (century time-windows)
0. Primary analysis	6.8	<.001	801	5.4	-6.6	<.001	200
1. Excluding religious hierarchy	6.8	<.001	801	5.3	-6.8	<.001	200
2a. Scale SC variables only	7.1	<.001	801	4.9	-6.1	<.001	200
2b. Non-scale SC variables only	6.5	<.001	801	5.9	-6.8	<.001	200
3. Moralizing high gods (MHG) only	6.8	<.001	801	-46.6	-8.7	<.001	159
4. With dating uncertainty	NA	NA	NA	4.4	-6.0	<.001	204
5a. 700-year max time-window	NA	NA	NA	5.6	-4.8	<.001	82
5b. Full time-windows	NA	NA	NA	3.9	-7.1	<.001	251
6a. Polity population only	6.8	<.001	801	5.3	-6.0	<.001	200
6b. Polity territory only	6.4	<.001	801	7.2	-6.0	<.001	200
6c. Capital population only	6.3	<.001	801	2.6	-6.5	<.001	200
6d. Hierarchical complexity only	6.9	<.001	801	6.5	-5.4	<.001	200
6e. Government only	5.5	<.001	801	5.9	-6.7	<.001	200
6f. Infrastructure only	5.4	<.001	801	7.0	-6.2	<.001	200
6g. Information systems only	5.9	<.001	801	9.8	-7.0	<.001	200
6h. Texts only	6.4	<.001	801	4.2	-5.4	<.001	200
6i. Money only	6.6	<.001	801	5.4	-4.8	<.001	200

See Methods for details, and see Extended Data Table 2 and Supplementary Tables 6–18 for full regression results. NA, not applicable.

Extended Data Table 5 | List of the 51 social complexity variables analysed

Variable no.	Name	Category	Sub-set
1	Polity population	Polity population	Scale
2	Polity territory	Polity territory	Scale
3	Population of the largest settlement	Population of the largest settlement	Scale
4	Administrative levels	Hierarchical complexity	Scale
5	Military levels	Hierarchical complexity	Scale
6	Religious levels	Hierarchical complexity	Scale
7	Settlement hierarchy	Hierarchical complexity	Scale
8	Professional military officers	Government	Non-scale
9	Professional soldiers	Government	Non-scale
10	Professional priesthood	Government	Non-scale
11	Full-time bureaucrats	Government	Non-scale
12	Examination system	Government	Non-scale
13	Merit promotion	Government	Non-scale
14	Specialized government buildings	Government	Non-scale
15	Courts	Government	Non-scale
16	Formal legal code	Government	Non-scale
17	Judges	Government	Non-scale
18	Professional Lawyers	Government	Non-scale
19	Irrigation systems	Infrastructure	Non-scale
20	Drinking water supply systems	Infrastructure	Non-scale
21	Markets	Infrastructure	Non-scale
22	Food storage sites	Infrastructure	Non-scale
23	Roads	Infrastructure	Non-scale
24	Bridges	Infrastructure	Non-scale
25	Canals	Infrastructure	Non-scale
26	Ports	Infrastructure	Non-scale
27	Mines or quarries	Infrastructure	Non-scale
28	Couriers	Infrastructure	Non-scale
29	Postal stations	Infrastructure	Non-scale
30	General postal service	Infrastructure	Non-scale
31	Mnemonic devices	Information system	Non-scale
32	Nonwritten records	Information system	Non-scale
33	Written records	Information system	Non-scale
34	Script	Information system	Non-scale
35	Non-phonetic writing	Information system	Non-scale
36	Phonetic alphabetic writing	Information system	Non-scale
37	Lists tables and classifications	Information system	Non-scale
38	Calendar	Texts	Non-scale
39	Sacred Texts	Texts	Non-scale
40	Religious literature	Texts	Non-scale
41	Practical literature	Texts	Non-scale
42	History	Texts	Non-scale
43	Philosophy	Texts	Non-scale
44	Scientific literature	Texts	Non-scale
45	Fiction	Texts	Non-scale
46	Articles	Money	Non-scale
47	Tokens	Money	Non-scale
48	Precious metals	Money	Non-scale
49	Foreign coins	Money	Non-scale
50	Indigenous coins	Money	Non-scale
51	Paper currency	Money	Non-scale

See a previous study[8] and http://seshatdatabank.info/methods/codebook/ for full definitions and selection rationale.

nature research

Corresponding author(s): Patrick E. Savage

Last updated by author(s): Feb 5, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All data was collected via, and is available open-access at: http://seshatdatabank.info/data/
Data analysis	All analyses were performed in R V3.4.2 and R Studio V 1.1.383. Source code is available online at http://github.com/pesavage/moralizing-gods

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The full machine-readable dataset is available as Supplementary Dataset 1, and at http://seshatdatabank.info/datasets. Full codings with detailed explanations and references are available at http://seshatdatabank.info/data, and are summarized in Extended Data Table 2. The data include the coded levels of uncertainty and disagreement, the textual explanations, and the references for each of the variables for all polities used in our analysis.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The global sample of 30 regions was selected and published (Turchin et al., 2015) before beginning data collection. The number of polities per region depended on the availability of historical and archaeological data, and was determined by the relevant expert historian/ archaeologist(s) for each region.
Data exclusions	The time period of interest was limited to between the beginning of the Neolithic to the beginning of Colonial/Industrial periods. No data falling within these periods was excluded.
Replication	No experiments were performed. Our primary results were replicated in all robustness analyses (see Methods).
Randomization	No randomization was performed.
Blinding	Historical data were coded by research assistants and regional experts who could not be blinded to the society they were coding. However, they were blind to the specific hypotheses being tested.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging